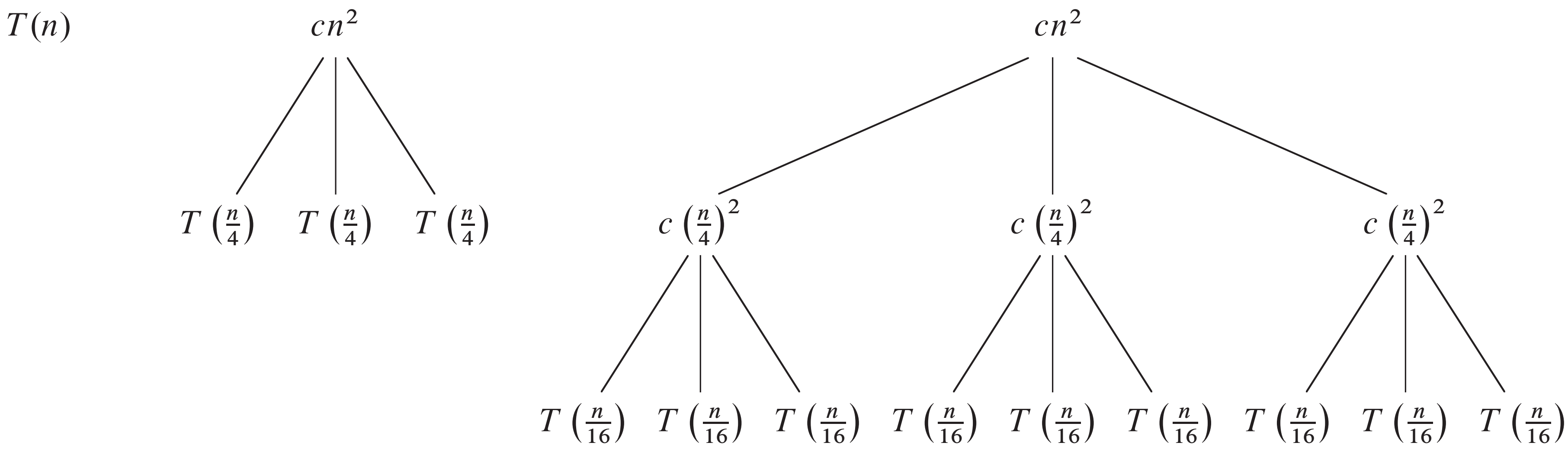
# Arbres de récursion

Approche alternative graphique pour *deviner* une solution analytique à une récurrence.

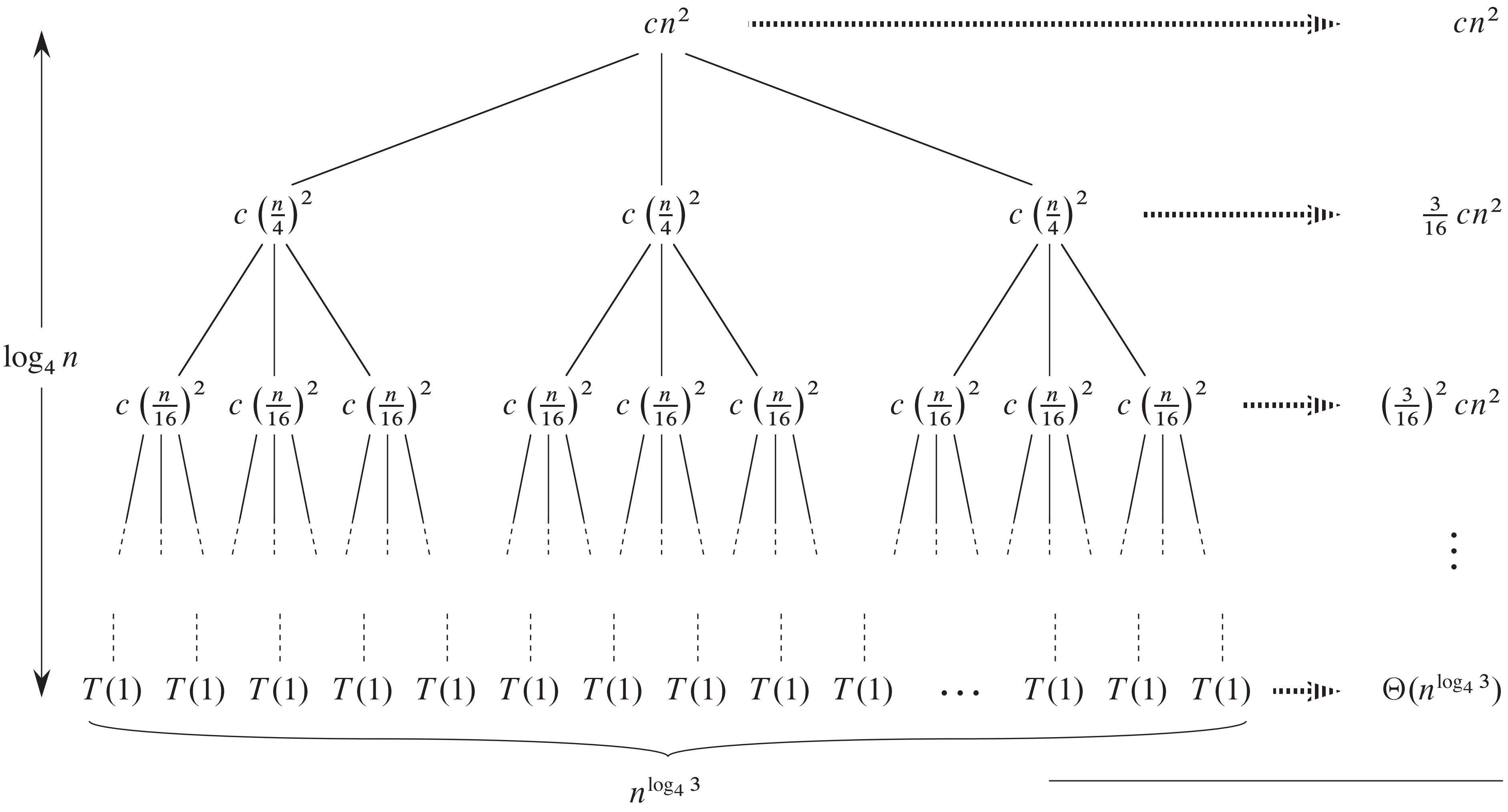
Illustration sur la récurrence suivante :

- $T(1) = a$
- $T(n) = 3T(n/4) + cn^2$ (Pour $n > 1$)

(Introduction to algorithms, Cormen et al.)



$T(n)$
$cn^2$
$T\left(\frac{n}{4}\right)$ $T\left(\frac{n}{4}\right)$ $T\left(\frac{n}{4}\right)$
$cn^2$
$c\left(\frac{n}{4}\right)^2$ $c\left(\frac{n}{4}\right)^2$ $c\left(\frac{n}{4}\right)^2$
$T\left(\frac{n}{16}\right)$ $T\left(\frac{n}{16}\right)$ $T\left(\frac{n}{16}\right)$ $T\left(\frac{n}{16}\right)$ $T\left(\frac{n}{16}\right)$ $T\left(\frac{n}{16}\right)$ $T\left(\frac{n}{16}\right)$ $T\left(\frac{n}{16}\right)$ $T\left(\frac{n}{16}\right)$


$cn^2$
$c\left(\frac{n}{4}\right)^2$ $c\left(\frac{n}{4}\right)^2$ $c\left(\frac{n}{4}\right)^2$
$\frac{3}{16}cn^2$
$\log_4 n$
$c\left(\frac{n}{16}\right)^2$
$\left(\frac{3}{16}\right)^2 cn^2$
$T(1)$ $T(1)$ $T(1)$ $T(1)$ $T(1)$ $T(1)$ $T(1)$ $T(1)$ $T(1)$ $T(1)$ … $T(1)$ $T(1)$ $T(1)$
$\Theta(n^{\log_4 3})$
$n^{\log_4 3}$

- Le coût total est la somme du coût de chaque niveau de l'arbre :

$$\begin{aligned}
T(n) &= cn^2 + \frac{3}{16}cn^2 + \ldots + \left(\frac{3}{16}\right)^{\log_4 n - 1} cn^2 + an^{\log_4 3} \\
&= \sum_{i=0}^{\log_4 n - 1} \left(\frac{3}{16}\right)^i cn^2 + an^{\log_4 3} \\
&< \sum_{i=0}^{\infty} \left(\frac{3}{16}\right)^i cn^2 + an^{\log_4 3} \\
&= \frac{1}{1-(3/16)} cn^2 + an^{\log_4 3} \\
&= O(n^2)
\end{aligned}$$

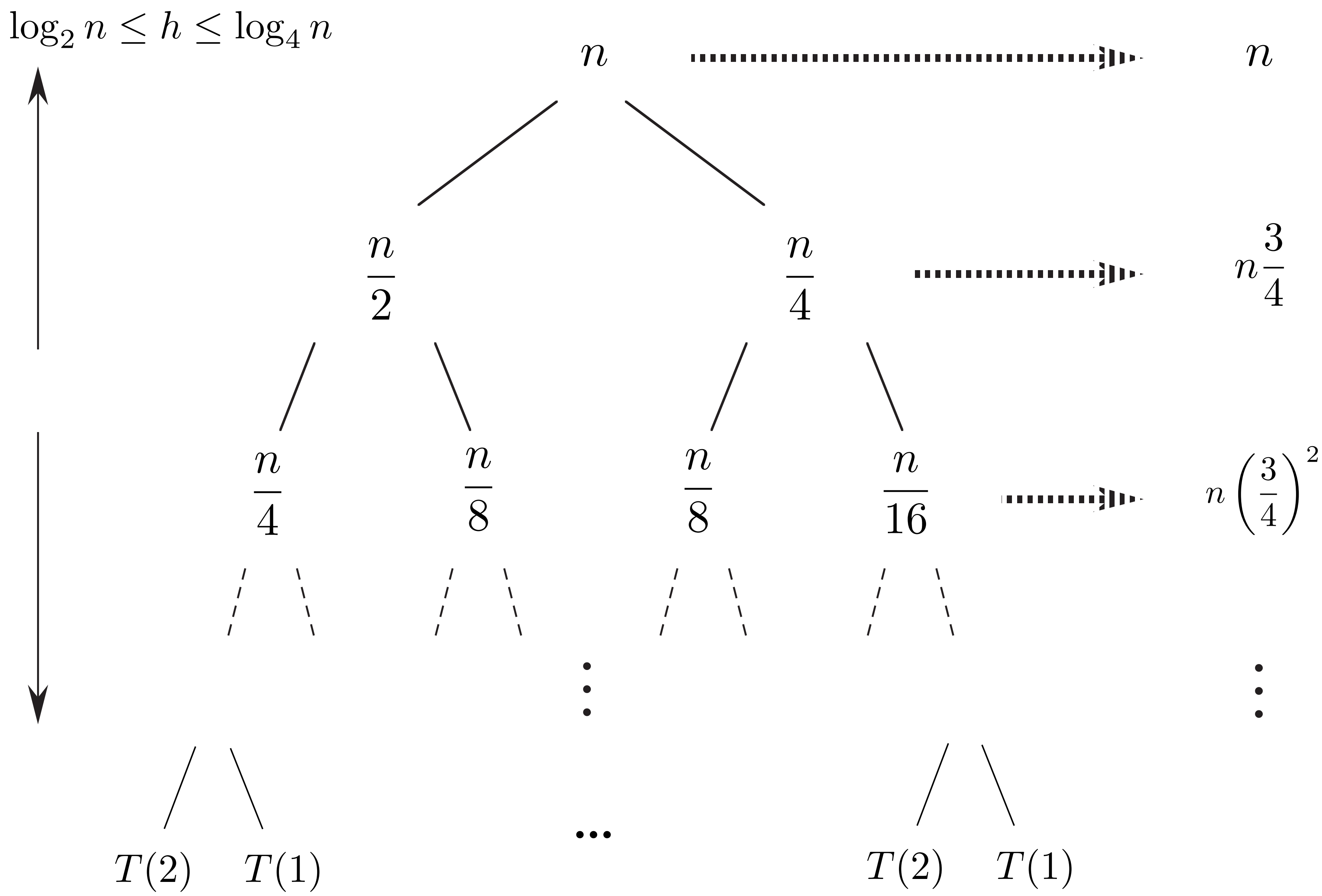
*(à vérifier par induction)*

- Comme le coût de la racine est $cn^2$, on a aussi $T(n) = \Omega(n^2)$ et donc $T(n) = \Theta(n^2)$.

Autre exemple :

- $T(1) = 1$
- $T(2) = 2$
- $T(n) = T(n/2) + T(n/4) + n$ (pour $n > 2$)

(On suppose que $n$ est toujours une puissance de 2)


$\log_2 n \le h \le \log_4 n$
$n$
$\frac{n}{2}$
$\frac{n}{4}$
$\frac{n}{4}$
$\frac{n}{8}$
$\frac{n}{8}$
$\frac{n}{16}$
$T(2)$ $T(1)$
$T(2)$ $T(1)$
$n$
$n\frac{3}{4}$
$n\left(\frac{3}{4}\right)^2$

- On déduit de l'arbre que

$$T(n) \leq \sum_{i=0}^{\infty} n \left(\frac{3}{4}\right)^i = n\frac{1}{1-\frac{3}{4}} = O(n)$$

- Vérification par induction forte qu'il existe un $c$ tel que pour tout $n \geq n_0$, on a $T(n) \leq cn$

$$\begin{aligned} T(n) &= T(n/2) + T(n/4) + n \\ &\leq cn/2 + cn/4 + n \\ &= (c3/4 + 1)n \\ &\leq cn \end{aligned}$$

$\Rightarrow$ ok pour tout $c > 4$

- Puisqu'on a aussi $T(n) = \Omega(n)$, on en déduit $T(n) = \Theta(n)$.

# Induction et notation asymptotique

Théorème faux : Soit la récurrence :

- $T(1) = 1$,
- $T(n) = 2T(n/2) + n$ (pour $n > 1$).

On a $T(n) = O(n)$.

*(la solution correcte est $T(n) = \Theta(n\log(n))$)*

Démonstration : (par induction forte)

- Soit $P(n) =$ "$T(n) = O(n)$".
- *Cas de base :* $P(1)$ est vrai car $T(1) = 1 = O(1)$.
- *Cas inductif :* Pour $n \geq 2$, supposons $P(1)$,$P(2)$,...,$P(n-1)$. On a :

$$\begin{aligned} T(n) &= 2T(n/2) + n \\ &= 2O(n/2) + n \\ &= O(n) \end{aligned}$$

*Où est l'erreur ?*

# Récurrences linéaires

Définition : Une *récurrence linéaire homogène* est une récurrence de la forme

$$f(n) = a_1 f(n-1) + a_2 f(n-2) + \cdots + a_d f(n-d)$$

où $a_1, a_2, \ldots, a_d \in \mathbb{R}$ sont des constantes. La valeur $d \in \mathbb{N}_0$ est appelée l'*ordre* de la récurrence.

Définition : Une *récurrence linéaire* (générale) est une récurrence de la forme

$$f(n) = a_1 f(n-1) + a_2 f(n-2) + \cdots + a_d f(n-d) + g(n),$$

où $g$ est une fonction ne dépendant pas de $f$.

Suite de ce chapitre : Étude d'une méthode permettant de *résoudre* les récurrences linéaires, c'est-à-dire de trouver des solutions analytiques équivalentes.

**Théorème :** Si $f_1(n)$ et $f_2(n)$ sont solutions d'une récurrence linéaire homogène (sans tenir compte des conditions initiales), alors toute combinaison linéaire $cf_1(n) + df_2(n)$ de $f_1(n)$ et $f_2(n)$ est également une solution pour tout $c, d \in \mathbb{R}$.

**Démonstration :** On a $f_1(n) = \sum_{i=1}^{d} a_i f_1(n-i)$ et $f_2(n) = \sum_{i=1}^{d} a_i f_2(n-i)$. Dès lors,

$$\begin{aligned} cf_1(n) + df_2(n) &= c \cdot \left( \sum_{i=1}^{d} a_i f_1(n-i) \right) + d \cdot \left( \sum_{i=1}^{d} a_i f_2(n-i) \right) \\ &= \sum_{i=1}^{d} a_i (cf_1(n-i) + df_2(n-i)). \end{aligned}$$

$\square$

# Exemple de résolution d'une récurrence

Supposons l'existence d'une nouvelle discipline scientifique, et des contraintes suivantes :

- $N$ places de professeurs enseignant cette discipline sont disponibles dans le monde.
- Chaque professeur
  - est nommé à vie ;
  - est supposé immortel ;
  - forme chaque année exactement un étudiant qui deviendra professeur l'année suivante (exception : lors de leur première année d'enseignement, les professeurs sont trop occupés pour former un étudiant).
- Année 0 : il n'y a aucun professeur.
- Année 1 : le premier professeur (autodidacte) est formé.

Question : Quand les $N$ places de professeurs seront-elles occupées ?

$\boxed{\text{Etape 1 : Trouver une récurrence}}$

| Année ($n$) | Nombre de professeurs ($f(n)$) |
|---|---|
| 0 | 0 |
| 1 | 1 (1 nouveau) |
| 2 | 1 (1 ancien qui forme un étudiant) |
| 3 | 2 (1 nouveau, 1 ancien) |
| 4 | 3 (1 nouveau, 2 anciens) |
| 5 | 5 (2 nouveaux, 3 anciens) |
| 6 | 8 (3 nouveaux, 5 anciens) |
| $\vdots$ | $\vdots$ |

Pour $n \geq 2$, on obtient $f(n) = f(n-1) + f(n-2)$.
Remarque : Il s'agit d'une récurrence linéaire homogène.

## Etape 2 : Résoudre la récurrence

- Une solution analytique pour une récurrence linéaire a souvent une forme exponentielle.
- On devine $f(n) = cx^n$ ($c$ et $x$ sont des paramètres à trouver).
-

$$
\begin{aligned}
& f(n) = f(n-1) + f(n-2) \\
\Rightarrow \quad & cx^n = cx^{n-1} + cx^{n-2} \\
\Rightarrow \quad & x^2 = x + 1 \\
\Rightarrow \quad & x = \frac{1 \pm \sqrt{5}}{2}
\end{aligned}
$$

- Les fonctions $c\left(\frac{1+\sqrt{5}}{2}\right)^n$ et $c\left(\frac{1-\sqrt{5}}{2}\right)^n$ sont des solutions de la récurrence (sans tenir compte des conditions initiales).
- Il en est de même pour toute combinaison linéaire de ces deux fonctions.
- On a donc $f(n) = c_1\left(\frac{1+\sqrt{5}}{2}\right)^n + c_2\left(\frac{1-\sqrt{5}}{2}\right)^n$.

- $f(0) = c_1 \left(\frac{1+\sqrt{5}}{2}\right)^0 + c_2 \left(\frac{1-\sqrt{5}}{2}\right)^0 = c_1 + c_2 = 0.$
- $f(1) = c_1 \left(\frac{1+\sqrt{5}}{2}\right)^1 + c_2 \left(\frac{1-\sqrt{5}}{2}\right)^1 = 1.$
- On obtient $c_1 = \frac{1}{\sqrt{5}}$ et $c_2 = \frac{-1}{\sqrt{5}}$.
- Finalement,

$$f(n) = \frac{1}{\sqrt{5}} \left(\frac{1+\sqrt{5}}{2}\right)^n - \frac{1}{\sqrt{5}} \left(\frac{1-\sqrt{5}}{2}\right)^n.$$

Réponse à la question : Les $N$ places de professeurs seront occupées lorsque $f(n)$ deviendra supérieur ou égal à $N$.

Comme $\left|\frac{1-\sqrt{5}}{2}\right| \approx 0,618 < 1$, cela se produira lorsque

$$f(n) \approx \frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^n \geq N.$$

Ce nombre d'années $n$ grandit donc logarithmiquement en fonction de $N$.

# Résolution des récurrences linéaires

Soit une récurrence de la forme

$$f(n) = a_1 f(n-1) + a_2 f(n-2) + \cdots + a_d f(n-d) + g(n)$$

et les conditions initiales $f(0) = b_0$, $f(1) = b_1$, etc.

| Etape 1 : Trouver les racines de l'*équation caractéristique* |
|---|

Définition : L'équation caractéristique est

$$x^d = a_1 x^{d-1} + a_2 x^{d-2} + \cdots + a_{d-1} x + a_d.$$

Remarque : Le terme $g(n)$ n'est pas pris en compte dans l'équation caractéristique.

Etape 2 : Trouver une *solution homogène*, sans tenir compte des conditions initiales

Il suffit d'additionner les termes suivants :

- Une racine non répétée $r$ de l'équation caractéristique génère le terme
$$c_r r^n,$$
où $c_r$ est une constante à déterminer plus tard.
- Une racine $r$ avec multiplicité $k$ de l'équation caractéristique génère les termes
$$c_{r_1} r^n, c_{r_2} n r^n, c_{r_3} n^2 r^n, \ldots, c_{r_k} n^{k-1} r^n,$$
où $c_{r_1}, c_{r_2}, \ldots, c_{r_k}$ sont des constantes à déterminer plus tard.

Etape 3 : Trouver une *solution particulière*, sans tenir compte des conditions initiales.

Une technique simple consiste à *deviner et vérifier* en essayant des solutions *ressemblant* à $g(n)$.

Exemples :

- Si $g(n)$ est un polynôme, essayer avec un polynôme de même degré, ensuite avec un polynôme de degré immédiatement supérieur, et ainsi de suite.
  Exemple : Si $g(n) = n$, essayer d'abord $f(n) = bn + c$, ensuite, $f(n) = an^2 + bn + c$, ...
- Si $g(n) = 3^n$, essayer d'abord $f(n) = c3^n$, ensuite $f(n) = bn3^n + c3^n$, $f(n) = an^2 3^n + bn3^n + c3^n$, ...

Remarque : On doit attribuer aux constantes $a, b, c, \ldots$ des valeurs satisfaisant l'équation récurrente.

Etape 4 : Former une *solution générale*, sans tenir compte des conditions initiales

Il suffit d'additionner la solution homogène et la solution particulière

Etape 5 : Déterminer les valeurs des constantes introduites à l'étape 2

- Pour chaque condition initiale, appliquer la solution générale à cette condition. On obtient une équation en fonction des constantes à déterminer.
- Résoudre le système formé par ces équations.

# Exemple

On demande de résoudre la récurrence suivante :

- $f(1) = 1$
- $f(n) = 4f(n-1) + 3^n$

$\boxed{\text{Etape 1 : Trouver les racines de l'équation caractéristique}}$

- L'équation caractéristique est $x = 4$.
- Sa seule racine est 4.

Etape 2 : Trouver une solution homogène, sans tenir compte des conditions initiales

La solution homogène est $f(n) = c4^n$.

Etape 3 : Trouver une solution particulière, sans tenir compte des conditions initiales.

- On devine que la solution est de la forme $d3^n$, où $d$ est une constante.
- En substituant, on obtient

$$\begin{aligned} d3^n &= 4d3^{n-1} + 3^n \\ 3d &= 4d + 3 \\ d &= -3 \end{aligned}$$

- On vérifie que $-3 \cdot 3^n = -3^{n+1}$ est bien une solution particulière.

Etape 4 : Former une solution générale, sans tenir compte des conditions initiales

On obtient la solution générale

$$f(n) = c4^n - 3^{n+1}.$$

Etape 5 : Déterminer les valeurs des constantes introduites à l'étape 2

$$\begin{aligned} f(1) = 1 \quad &\Rightarrow \quad c4^1 - 3^{1+1} = 1 \\ &\Rightarrow \quad c = \frac{5}{2}. \end{aligned}$$

Conclusion : $f(n) = \dfrac{5}{2}4^n - 3^{n+1}$.

# Changement de variables

Un changement de variables permet parfois de réduire une récurrence "diviser pour régner" à une récurrence linéaire.

Soit la récurrence du transparent 241 :

$$T(n) = 7T(n/2) + O(n^2)$$

En posant : $n = 2^m$ et $S(m) = T(2^m)$, on obtient :

$$S(m) = T(2^m) = 7T(2^{m-1}) + O((2^m)^2) = 7S(m-1) + O(4^m)$$

# Résumé

- Outils de résolution d'équations récurrentes :
  - Méthodes génériques : "Deviner-et-Vérifier", "Plug-and-Chug", arbres de récursion
  - Récurrences "Diviser pour régner" : théorème "Master", Théorème d'Akra-Bazzi
  - Récurrences "linéaires"

  Les deux dernières sont les plus systématiques.
- Le plus dur reste de traduire un problème réel en une équation récurrente.
- Exemple : Soit un type de plante qui vit éternellement, mais qui peut seulement se reproduire la première année. A quelle vitesse la population croît-elle ?

# Comparaison

| | Récurrence | Solution |
|---|---|---|
| Tours de Hanoï | $T_n = 2T_{n-1} + 1$ | $T_n \sim 2^n$ |
| Tours de Hanoï 2 | $T_n = 2T_{n-1} + n$ | $T_n \sim 2 \cdot 2^n$ |
| Algo rapide | $T_n = 2T_{n/2} + 1$ | $T_n \sim n$ |
| Tri par fusion | $T_n = 2T_{n/2} + n - 1$ | $T_n \sim n \log n$ |
| Fibonacci | $T_n = T_{n-1} + T_{n-2}$ | $T_n \sim (1.618\ldots)^{n+1}/\sqrt{5}$ |

- Récurrences "Diviser pour régner" généralement polynomiales
- Récurrences linéaires généralement exponentielles